Final Project Proposal: Sentiment Analysis on Amazon Product Reviews

Course: CS6120 Natural Language Processing
Academic Year: Fall 2023
Instructor: Dr. Kenneth Church
Teaching Assistant: Ferran Sulaiman
Student Name: **Jiali Han**
Student Email: han.jial@northeastern.edu

Project Objective

To design and develop a sentiment analysis model that can effectively analyze and classify the sentiments of Amazon product reviews.

Project Introduction

My project aims to perform sentiment analysis on Amazon product reviews using the **NLTK (Natural Language Toolkit)** and **Transformers**, a state-of-the-art deep learning library. By identifying the sentiment of these reviews, we can gain a better understanding of customer opinions, product satisfaction, and market trends.

Why is it worth doing?

Feedback for Sellers and Manufacturers: By identifying the sentiment behind customer reviews, sellers and manufacturers can understand their products' strengths and weaknesses, leading to product improvement.

Improved Decision Making for Buyers: A summary of sentiments can guide potential buyers in understanding the general reception of a product without having to go through each individual review.

Business Strategies: Sentiment analysis can uncover trends and shifts in consumer preferences. By analyzing patterns in positive or negative reviews across different products, one can compare how products from different brands are received by the public. In this way, companies can also tweak their marketing strategies and product development strategies based on the positive or negative sentiment trends.

Research Possibilities: This project provides an opportunity to explore and apply both traditional NLP techniques (using NLTK) and modern deep learning approaches (using Transformers) for sentiment analysis. Since this is my first time learning NLP, this project will be a valuable educational experience and research opportunity for me and also for those interested in NLP and machine learning.

Feasibility

Data Availability: Amazon product reviews are readily available. With the right web scraping tools and techniques, I believe I am able to collect a substantial dataset for analysis.

NLP Tools: NLTK provides a comprehensive set of tools for NLP tasks, including tokenization, stemming, and sentiment analysis, which can be used for the initial analysis. Transformers, on the other hand, offers pre-trained models that can achieve state-of-the-art results in sentiment analysis tasks.

Community and Resources: NLTK and Transformers have strong communities and extensive documentation, making it feasible to find support, tutorials, and code examples to help you get started.

Computational Resources: Although deep learning models require significant computational power, cloud platforms like AWS, Google Cloud, and Azure provide necessary GPU and TPU resources for model training and deployment. Therefore, I'm planning to use Google Colab as the Cloud platform and service that helps me access resources for research and development.

In summary, my project aims to leverage the NLP techniques and provide a reliable sentiment analysis model for Amazon product reviews, so it can provide insights into consumer preferences, trends in the market, and the opportunity to apply both traditional and cutting-edge NLP techniques in real-world scenarios.